



Nurse Effects on Non-response in Survey-Based Biomeasures

Alexandru Cernat , Joseph W. Sakshaug , Tarani Chandola , James Nazroo & Natalie Shlomo

To cite this article: Alexandru Cernat , Joseph W. Sakshaug , Tarani Chandola , James Nazroo & Natalie Shlomo (2020): Nurse Effects on Non-response in Survey-Based Biomeasures, International Journal of Social Research Methodology, DOI: [10.1080/13645579.2020.1832737](https://doi.org/10.1080/13645579.2020.1832737)

To link to this article: <https://doi.org/10.1080/13645579.2020.1832737>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 15 Oct 2020.



[Submit your article to this journal](#)



Article views: 214



[View related articles](#)



[View Crossmark data](#)

Nurse Effects on Non-response in Survey-Based Biomeasures

Alexandru Cernat^a, Joseph W. Sakshaug^b, Tarani Chandola^a, James Nazroo^a
and Natalie Shlomo^a

^aUniversity of Manchester, Manchester, UK; ^bInstitute for Employment Research (IAB), Ludwig Maximilian University of Munich, University of Mannheim, Mannheim, Germany

ABSTRACT

Collecting biological data in representative surveys is becoming more common due to their potential to inform research and policy. Nevertheless, using nurses to collect these data can lead to unintended effects. In this paper, we investigate how nurses influence the non-response process by looking at five waves of data coming from two surveys in the UK: Understanding Society and the English Longitudinal Study of Ageing. We find that nurses explain between 5 and 14% of the variance in non-response to biological data collection. We also find that older nurses are more successful in obtaining cooperation and consent to biological data collection and nurses with more survey experience are more likely to successfully collect blood samples. Finally, we show that including nurse characteristics in non-response weighting models leads to modest changes in population estimates of biological markers.

KEYWORDS

Non-response; bio-social research; longitudinal data; weighting; nurse effects

1. Introduction

There is increasing interest in combining social data with objective physical and biological measurements (hereafter referred to as ‘biomeasures’) in surveys (Sakshaug et al., 2015). These so-called ‘bio-social surveys’ allow researchers to obtain a deeper understanding of the interplay between genetics, environmental factors, and behaviors and the role these processes play in determining a variety of social- and health-related outcomes (National Research Council (NRC), 2001, 2008). Household surveys that collect self-report data commonly supplement these data with an array of biomeasures, ranging from anthropometric measures (e.g. height/weight, waist circumference) and physical performance assessments (e.g. walking speed, grip strength) to biological specimens (e.g. blood, saliva). Such measures are collected in several large-scale surveys, such as the U.S. Health and Retirement Study (HRS), the Survey of Health, Ageing and Retirement in Europe (SHARE), and the English Longitudinal Study of Ageing (ELSA). Collecting biomeasures and analyzing them in conjunction with social data has led to discoveries in multiple fields of inquiry, such as cognitive aging (e.g. Seeman et al., 2005) and physical functioning (e.g. Elosua et al., 2005).

While incorporating biomeasures into household surveys increases research opportunities, it also introduces methodological issues. One issue is non-response. There are at least three possible ways in which biomeasure non-response can occur. For instance, not all respondents are willing to participate in the biomeasure component of a survey (Banks et al., 2006; Kearney et al., 2011; S. McFall et al., 2014). Further, even when cooperation is obtained, respondents may not consent to every individual biomeasure. The general issue of non-consent to health survey data collection has garnered much attention in the literature, including in this very journal (Al Baghal, 2016; Mostafa,

CONTACT Alexandru Cernat  Alexandru.cernat@manchester.ac.uk  University of Manchester, Manchester M13 9PL, UK
 Supplemental data for this article can be accessed [here](#).

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

2016; Mostafa & Wiggins, 2018). Non-consent is particularly common for blood collection – one of the more invasive survey biometrics (Jaszczak et al., 2009; S.L. McFall et al., 2014; Sakshaug et al., 2010; Weiss et al., 2019). The last stage of non-response is the failure to obtain a valid measurement from a willing participant. This failure may be caused by physical inability (e.g. diagnosed blood clot), equipment malfunction, or administrator error.

Depending on the costs and goals of the study, medically trained personnel, such as nurses, or non-medically trained lay interviewers are used to collect biometrics in household surveys. Both actors play a critical role in the three stages of biometric participation. Lay interviewers are known to vary in their ability to successfully collect biometrics from survey respondents (Jaszczak et al., 2009; Korbmacher, 2014; Sakshaug et al., 2015). Further, interviewer characteristics (e.g. demographics, experience) can influence respondents' likelihood of biometric participation (Korbmacher, 2014; Sakshaug et al., 2015). Such interviewer effects have the potential to adversely affect analyses of biometric data if they are not taken into account.

Little is known regarding the effects of nurses on respondent participation in the biometric component of a survey. While nurses are more qualified than lay interviewers to collect biometrics, they typically have less experience working in survey settings and gaining respondent cooperation in households. In this study, we examine nurse effects in two nationally-representative bio-social surveys that collected whole blood in respondents' homes. We make use of detailed nurse data to assess the influence of nurses on the likelihood of obtaining cooperation, consent, and collecting a valid whole blood sample from respondents. The remainder of this article is organized as follows. In [Section 2](#) we review the biometric literature with a particular focus on the effects of interviewers and nurses. In [Section 3](#) we describe the data sources and analytic approaches we use to study nurse effects. The results of the analysis are presented in [Section 4](#). A general discussion of the findings and practical implications for household biometric collection are provided in [Section 5](#).

2. Background

There are different approaches that household surveys use to collect biometrics from respondents. The approaches can be divided into clinic-based and home-based collection. Clinic-based approaches examine respondents in a clinical setting, e.g. at a hospital or medical center where respondents must present. One example is The Irish Longitudinal Study of Ageing, which arranges examinations performed by trained medical personnel at a clinical site (with the option of a home assessment for respondents unable to travel) (Kenny et al., 2010). A variation of this approach examines respondents in a mobile clinical setting. For example, the U.S. National Health and Nutrition Examination Survey carries out physical assessments by trained medical personnel inside mobile examination centers (Zipf et al., 2013). These clinic-based approaches offer sophisticated biometric equipment, storage, and laboratory capabilities, but at a very high cost.

Home-based biometric collection is a comparatively less expensive option used in many countries. There are different approaches of home-based biometric collection that utilize different personnel. A common approach is to use non-medically trained interviewers to collect minimally invasive biometrics. For example, SHARE uses lay interviewers to collect biometrics in multiple countries (Weiss et al., 2019). A key advantage of this approach is the completion of both the main interview and biometric component in a single visit. This approach also accommodates respondents unable to travel to a medical facility, who tend to have different health profiles compared to able respondents (Kearney et al., 2011). On the other hand, using lay interviewers restricts the range of biometrics that can be collected. In the UK, for instance, lay interviewers are permitted to collect dried blood spots using minimally invasive finger prick methods, but not more invasive techniques, such as intravenous blood collection.

Another home-based approach – and the focus of our study – uses trained medical professionals, or nurses, to carry out the biomeasure collection in respondents' homes. This is the approach used by ELSA and Understanding Society – The UK Household Longitudinal Study (UKHLS). Both studies deploy nurses to visit respondents after the main interview. This approach allows for a wider range of biomeasures to be collected (e.g. whole blood). However, the gap between the main interview and nurse visit can be rather long (e.g. several months), giving respondents who ostensibly agree to participate ample time to reconsider their decision.

Biomeasure cooperation is not universal and consent rates to individual biomeasures can vary widely (Jaszczak et al., 2009; S.L. McFall et al., 2014; Sakshaug et al., 2010). Nurses and interviewers alike play an important role in obtaining biomeasure cooperation from respondents. Just as interviewers vary in their ability to recruit survey participants and collect medical information – a history that dates back to the 1950's (e.g. Cochrane et al., 1951; Durbin & Stuart, 1951) – nurses are likely to vary in their ability to convince respondents to participate in biomeasure collection and obtain valid measurements from them. Administrator effects in survey-based biomeasure collections are understudied and the literature tends to focus on the effects of interviewers rather than nurses. Jaszczak et al. (2009) found interviewer variability in biomeasure cooperation rates in the National Social Life, Health, and Aging Project (NSHAP), with larger variability for the most invasive biomeasures. For example, they report interviewer cooperation rates between 75 and 100 percent for weight measurements, and between 0 and 100 percent for the collection of a self-administered vaginal swab. Sakshaug et al. (2010) found a statistically significant between-interviewer variance for consent to blood collection in the HRS. Korbmacher (2014) also reported statistically significant interviewer variation in consent to collect blood in the SHARE study. It is worth noting that these studies did not randomly assign interviewers to respondents; thus, variation in biomeasure cooperation rates could be (at least, partially) explained by interviewer allocation to more difficult respondents or areas where contacting and obtaining cooperation from respondents is more challenging (e.g. urban areas).

Another potential source of interviewer variation is due to characteristics of the interviewer, which have a long history of inquiry (e.g. Steinkamp, 1966; for a review, see West & Blom, 2017). Korbmacher (2014) showed, for example, that interviewers' age (+), education (+), survey experience (-), expectations regarding biomeasure success (+), among other characteristics, were significant predictors of obtaining respondent consent to blood collection. Accounting for interviewer characteristics considerably reduced the proportion of total variation explained by interviewers (from 36 to 9%). Williams and McDade (2009) also report a positive effect of interviewer experience on blood collection in the NSHAP study. Lindau et al. (2009) found that interviewers whose age matched that of the NSHAP study population (ages 57–85) achieved higher cooperation rates for the vaginal self-swab protocol. Sakshaug et al. (2010) found no effect of interviewer demographics and experience on respondents' likelihood of biomeasure consent with the exception of race/ethnicity: African-American interviewers were less likely to obtain consent from respondents compared to interviewers from other racial groups.

Compared to the above interviewer studies, the effects of nurses on biomeasure participation is far less established. To our knowledge, only one published study has examined nurse effects in survey-based biomeasures collected in respondents' homes. Anglewicz (2009) sent nurses to the homes of participants of the Malawi Diffusing and Ideational Change Project 2004–2006, a study which took place in three sites in Rural Malawi. A significant gender interaction between respondents and nurses was found for HIV testing, in which male respondents from the southern site who were visited by a male nurse were more likely to take part in HIV testing compared to male respondents visited by a female nurse.

Understanding the impact of nurses at multiple stages of biomeasure participation (e.g. nurse visit cooperation, consent to individual biomeasure, etc.) would be useful for researchers considering whether to adopt the nurse setting in bio-social surveys. We address this topic by analyzing nurse effects and the variation explained by nurses on whole blood collection in two nationally-

representative bio-social surveys in the UK – the English Longitudinal Study of Ageing and Understanding Society – The UK Household Longitudinal Study. We use detailed information on nurses supplemented with geographical information to study the effects of nurses on biomeasure participation while disentangling these effects from the areas they work in. Specifically, we answer the following research questions:

- (1) How much variation is explained by nurses in the likelihood of participation at each stage of the biomeasure component of a bio-social survey?
- (2) What are the individual-level predictors that explain participation in the nurse visit and blood collection?
- (3) Do characteristics of nurses influence the likelihood of respondents participating in the nurse visit and blood collection?
- (4) Does adjusting for nurse characteristics impact population estimates of biological markers?

3. Data and methods

Data source

The UK Household Longitudinal Study (University of Essex, 2018) is a yearly household survey representative of the United Kingdom population. The latest re-adaptation of the survey, known as Understanding Society, started collecting data from 2008. From wave 2 onward, the British Household Sample (BHPS) was also included in the data collection. The survey collects data from all household members 16 years or older while children between 10 and 15 years old are invited to take a self-completion youth questionnaire. In wave 1 of the UKHLS the response rates were 57.3% at the household level and 81.8% at the individual level. In wave 2 the household response rate was 61.7% while the individual rate was 59.4%. Finally, in wave 3 response rates were 57.3% and 61.3% for households and individuals, respectively (Knies, 2018).

In wave 2 of the survey a random subsample (81% of Primary Sampling Units) of the General Population Sample was selected for a nurse visit (University of Essex, 2014) due to funding limitations. We refer to this as the Understanding Society Wave 2 sample (USW2). In wave 3, eligible BHPS sample respondents were invited for a nurse visit. We refer to this sample as the BHPS Wave 19 sample (BHPSW19). Waves 2 and 3 were conducted between 2010 and 2012 by NatCen Social Research.

Nurse visits were conducted by professional nurses in respondents' homes approximately 6 months after the main interview. Before the nurse visit, eligible respondents received an advance letter and leaflet explaining the nurse visit. The eligibility criteria were: completion of the face-to-face interview in the respective wave, aged 16 or older, living in England, Scotland, or Wales, completing the main interview in English, and not being pregnant (McFall et al., 2014). Initial contact was made by telephone staff who attempted to arrange an appointment. An attempted nurse visit did not always lead to participation as non-contacts and refusals were still possible. During the nurse visit, oral consent was needed for anthropometric measures and written consent was needed for blood samples. Measurements were recorded using Computer Assisted Personal Interviewing (CAPI). Overall response rates for the nurse visit (out of those eligible) were 58.6% for USW2 and 57% for BHPSW19 while blood was collected from 38.2% for USW2 and 37.5% for BHPSW19 of the eligible sample (McFall et al., 2014).

The English Longitudinal Study of Ageing is a longitudinal study aiming to investigate aging for individuals 50 years or older (Blake et al., 2018). The study started in 2002 and the original sample was based on respondents to the Health Survey for England (HSE) conducted between 1998 and 2001. The sample was refreshed at waves 3, 4, 6, and 7 with new HSE respondents. Data were collected every 2 years (NatCen Social Research, 2015). Response rates were 70% at the household level and 67% at the individual level in wave 1. In waves 2, 4, and 6 the individual response rates

were 82%, 71%, and 76%, respectively, of those eligible (Bridges et al., 2015; Cheshire et al., 2012; Scholes et al., 2008; Taylor et al., 2007). The exact numbers of respondents at each stage in all five datasets are presented in Figure A1 of the online appendix. These are the cases used in the analysis and excludes cases with missing data, which we discuss in the next section.

In waves 2, 4, and 6 eligible respondents were invited to participate in the nurse visit. We refer to these samples as ELSAW2, ELSAW4, and ELSAW6. Eligible respondents were those who: were original sample members in ELSA (core members), had an in-person interview in that wave, didn't have clotting/bleeding, and had no history of fits or were not taking anticoagulant drugs. Response rates for the nurse visits in waves 2, 4, and 6 were 87.3%, 85.7%, and 84.3%, respectively, among those eligible (NatCen Social Research, 2015). Similar to the procedures used in UKHLS, the visit was conducted by professional nurses using CAPI at the respondent's home and written consent was required for blood collection.

In this paper, we investigate non-response in the nurse visit in five waves of data collection: USW2, BHPSW19, ELSAW2, ELSAW4, and ELSAW6. We distinguish between three stages of non-response. The first one is the cooperation to have a nurse visit conditional on being eligible. The second one is consent to blood collection conditional on nurse visit participation. Finally, we examine the likelihood of successfully collecting blood given consent.

There are two reasons why we separate these stages. Firstly, we expect differences in the impact of nurses on participation at each stage. In the first stage, we expect smaller nurse effects due to failed appointment attempts made by the telephone staff or doorstep refusals. Further, the types of skills needed to convince people to consent to give blood and actually collecting valid blood samples may be distinct. The second reason we separate these stages is because we are interested in possible explanations for non-response. We expect the reasons people do not participate at each stage to be different. Thus, we expect the influences of first-stage non-response to be similar to non-response in surveys in general, the second stage to be more similar to other types of consent where privacy and data confidentiality come into play (e.g. data linkage consent), while the last stage might be mainly explained by health reasons and nurse skills. A better understanding of these mechanisms can lead to better correction for missing data (e.g. through modeling or weighting/imputation approaches) and better data collection approaches (e.g. through improved training and intervention).

The independent variables used in the analysis are: gender, having a partner, living alone, being white, age, education, owning a house, living in London, living in the North, self-reported health, having a long-term illness, nurse race, nurse age, and nurse experience. In UKHLS we include additional variables: in work, urban, very interested in politics. These variables were not included in ELSA either because they had little variation (e.g. fewer older people work) or the variables were not available in the database. A nurse gender variable (which was predominantly female) did not have enough variation to warrant inclusion in the models. The independent variables were selected based on supporting literature which finds that demographics (Tolonen et al., 2006), health status (Cohen & Duffy, 2002), and interest in politics (Couper, 1997) to be strong predictors of survey participation. The distribution of the variables for all five waves can be found in Tables A1 and A2 of the online appendix. We do not investigate the implications of nurse workloads and nurse continuity on longitudinal nonresponse, as this topic has been addressed in separate work (Cernat & Sakshaug, 2020).

Statistical methods

The statistical approach used here closely mimics that used in the interviewer effects literature (West & Blom, 2017). The main difficulty in estimating the impact of nurses on participation stems from the fact that they are not randomly allocated to respondents. The main potential confounders are related to area effects (e.g. urban areas tend to have lower participation rates in surveys, wealthier areas may be visited by nurses with better training opportunities). One way to partially separate these two effects is to estimate cross-classified models (Snijders & Bosker, 2011; West &

Blom, 2017). In these models, we decompose the total variation of individual participation into nurse effects, area effects, and residual variance. The mathematical notation is:

$$Y_{i(j,k)} = \gamma_0 + \sum^{y_h} x_{i(j,k)} + U_{0j} + U_{0k}$$

where the dependent variable Y varies by individual (i), area (j) and nurse (k). The fixed part of the model has an intercept (γ_0) and h control variables with fixed effects (γ_h). We also specify random effects for area (U_{0j}) and nurse (U_{0k}). Because the dependent variables are dichotomous, we estimate a probit regression (for more details see online appendix).

For area effects, we use the Lower Super Output Area (LSOA) which is a standard area measure proposed by the Office for National Statistics and includes approximately 1,500 households. For nurse effects, we use nurse IDs as provided by the data collection agency.

To estimate the models, we use Bayesian estimation as implemented in Mplus 8 (Muthén & Muthén, 2017). This modeling approach is preferred as it can better handle complex models when there are borderline coefficients (coefficients with random effects close to 0). More details on the estimation can be found in the online appendix. We use listwise deletion for missing data. The proportion of missing data is around 2.6% (or 234 cases) over all ELSA waves. For USW2 there is around 18% missing data (4,795 cases). The main reason for the missing data is the missing nurse IDs for 4,213 cases (15%). Almost all these cases (4,210) were for respondents who did not participate in the nurse visit. Unfortunately, this information could not be recovered. Nurse race also had considerable missing data in USW2 (around 3,000 cases or 11.3%) and as such we added a category if the nurse information is missing or not in the model. For BHPSW19, there is around 5% missing data (or 306 cases).

For each wave of data collection and each outcome we run three models:

- Model 0: an empty model explaining the outcome using only the intercept and random effects for areas and nurses.
- Model 1: Model 0 + individual characteristics from the main survey.
- Model 2: Model 1 + nurse characteristics as provided by the data collection agency.

Correcting for missing data using weighting

For the last research question, we investigate the use of nurse characteristics in adjusting for non-response and whether this impacts population estimates of biological markers, namely, average pulse rate, and C-Reactive Protein (CRP). To do this, we compare estimates from the nurse visit (first stage) and blood collection (third stage) using different weighting approaches. To highlight the exact effect of including nurse characteristics in non-response correction models we develop a series of nested weights:

- (1) Interview weights from the survey organization and area clustering
- (2) + individual characteristics predicting each stage of non-response (based on model 1)
- (3) + nurse characteristics predicting each stage of non-response (based on model 2)
- (4) + adding nurse clustering (based on nurse ID)

From the nurse visit, we estimate the average pulse rate in the population. For the pulse rate variable, we construct the weight by multiplying the interview and nurse visit weights:

$$weight_{pulse} = weight_{interview} * weight_{nursevisit}$$

From the blood collection, we estimate mean CRP in the population. This is an indicator of stress and is often used in health research. To calculate the appropriate weight for the CRP we multiply the weights at all four stages:

$$weight_{CRP} = weight_{interview} * weight_{nursevisit} * weight_{consent} * weight_{blood}$$

The interview weight is provided by the survey agency. The other three weights are based on the inverse of the probability of responding at that particular wave. The response probabilities are calculated using the fixed effects from models 1 (for the second weight) and model 2 (for the third and fourth weights). Following suggestions from Rosenbaum and Rubin (1983), we performed a propensity score subclassification to minimize weight variation. The procedure was implemented by sorting the propensities from low-to-high and creating five roughly equal-sized propensity score groups. The average response propensity within each quantile was then retained and the weight was constructed based on the inverse of this value. We use R 3.5.1 for all data cleaning and visualization, and the *survey* package for all weighting (Lumley, 2004).

4. Results

Estimation of nurse effects

To answer the first research question, what is the impact of nurses on the likelihood to participate, the total variation of the response process is decomposed into three components: nurse effects, area effects, and residual. This is done for all three models and all five datasets.

Figures 1 and 2 show the results of this decomposition. Overall, nurses have small-to-medium-sized effects that are smaller for the first-stage nurse visit participation, and become slightly larger for consent to blood collection, and actual blood collection. The average proportion of variance across all models for UKHLS is 7% for nurse visit, 8% for blood consent, and 12% for blood collection. For ELSA, the results are similar with the average proportion of variance being 5% for the nurse visit, 10% for blood consent, and 14% for blood collection.

The two figures show that the proportion of variance is relatively constant between statistical models even after controlling for individual and nurse characteristics. The nurse effects are also similar within the waves of the two surveys. USW2 and BHPSW19 on the one hand and ELSAW2, ELSAW4, and ELSAW6 on the other hand have relatively consistent nurse effects. One noticeable difference is the reduced area effects in the nurse visit component. ELSAW2 has a large area effect that seems to be decreasing in time over the subsequent waves. The reduced area effect over time could be due to the study having a greater proportion of more engaged/committed respondents and thus the reasons for non-response will have more to do with individual circumstances than with

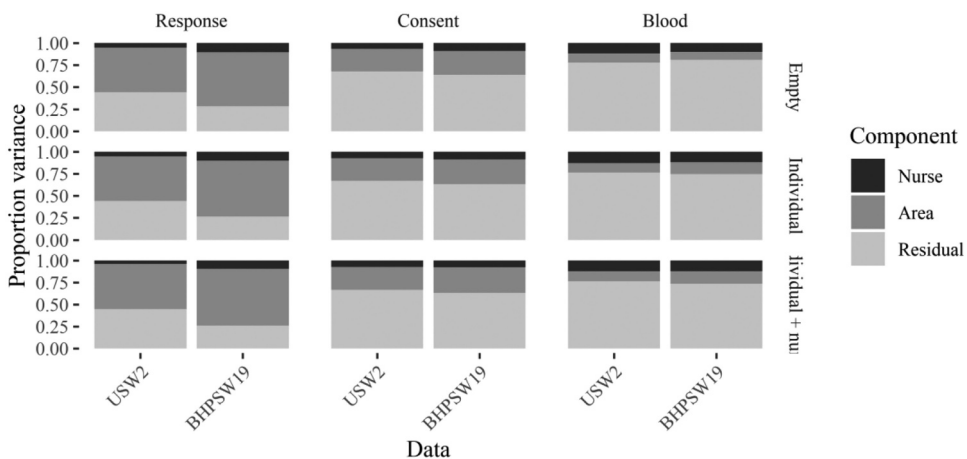


Figure 1. UKHLS variance decomposition for the three non-response stages by data and statistical model.

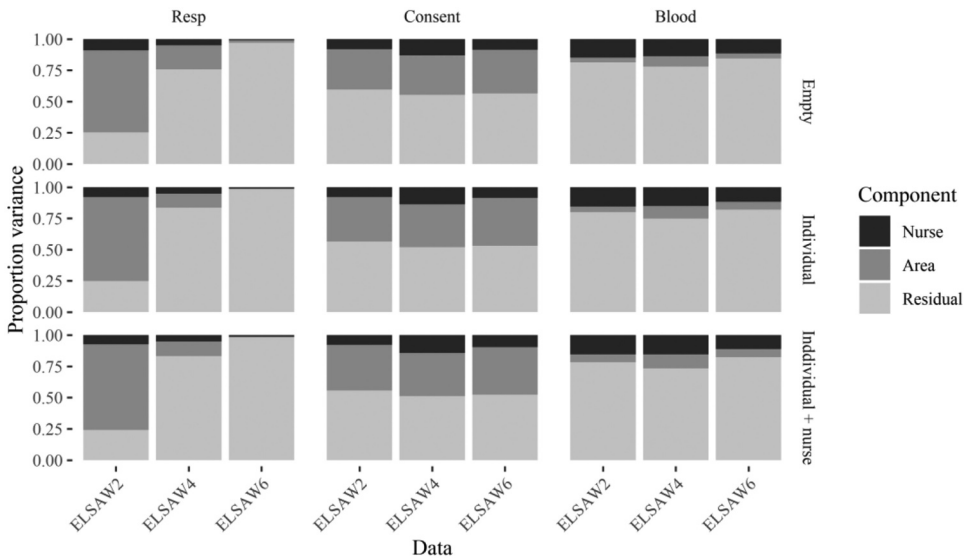


Figure 2. ELSA variance decomposition for the three non-response stages by data and statistical model.

anything related to area (e.g. social class). This was a potential explanation offered by NatCen when this finding was brought to their attention (personal communication, 26 October 2018).

Impact of individual characteristics on participation

For the next research question, we look at the effects of individual characteristics on the likelihood to participate in the three stages. The results from model 3 are reported as this also controls for nurse characteristics. Figure A2 and A3 show the probit coefficients with the credibility intervals by response stage and dataset (see Tables A3–A8 for exact coefficients).

For UKHLS, older respondents are more likely to participate in all three stages. This is true for both datasets. People with lower education tend to be less likely to participate in the nurse visit compared to those with a degree. Women are more likely to participate in the nurse visit but less likely to collect blood from. Having a partner increases the chances of participation in all three stages of USW2 but has a negative effect on actual blood collection in BHPSW19. Larger households appear less likely to participate in the nurse visit in both waves. Living alone, on the other hand, increases the likelihood of nurse visit participation and consenting to blood in USW2, and decreases the likelihood of collecting blood in BHPSW19. Being white increases the chances of participating in all three stages in USW2. Owning a home increases the chances of participating in the nurse visit in both waves, and increases the chances of consenting and collecting blood for BHPSW19. Interest in politics increases the chances of nurse visit participation in both surveys while living in London decreases the chances in USW2. Having a long-term illness increases the chances of participating in the nurse visit and consenting to blood for USW2. Higher (better) self-reported health increases the chances to participate in all three stages of USW2 and increases the chances to collect blood for BHPSW19.

Looking at the effect sizes, age is one of the strongest predictors of participation. Respondents over 56 years of age are over 20% more likely to participate in the nurse visit compared to those under 35 in both surveys. Similarly, those between 35 and 56 years-old are 14% more likely to participate in the nurse visit compared to younger respondents. These effects are also present for consent to blood collection. White respondents are 16% more likely to consent to give blood in USW2 compared to non-whites and 6% more likely to actually give blood compared to those that are non-white.

Looking at results from ELSA, older respondents are more likely to participate in the nurse visit but are less likely to consent to give blood. This is consistent in all three waves. In ELSAW6, older respondents are less likely to give blood (conditional on consent). People with no qualifications are less likely to participate in the nurse visit in wave 2 while females are more likely. Similar to UKHLS, females are also less likely to give blood in all three waves of ELSA. Having a partner decreases the chances to participate in the waves 4 and 6 nurse visits. This is also true for living alone, which decreases the chances to participate in the nurse visit and consent to blood but increases the chances of collecting blood in wave 4. Living alone also has a detrimental effect on nurse visit participation in wave 6. Respondents in the North are less likely to respond in waves 2 and 4 while those in London are less likely to give blood in wave 6. Owning a house increases the chances to participate in the nurse visit and consent to collecting blood in waves 2 and 6 while being white increases the chances to participate in the nurse visit in wave 2 by 23% (compared to non-white) and consent in waves 4 and 6 by 17% and 11%, respectively. Having a long-standing illness increases the chances to respond to the nurse visit in waves 2 and 4 while higher self-rated health increases the chances to participate in all stages and all waves of ELSA.

Impact of nurse characteristics on nurse participation

Next we examine the effects of nurse characteristics on participation. For UKHLS, older nurses are more likely to obtain nurse visit participation compared to younger ones both in USW2 and BHPSW19. If they are over 64 years-old they are 17% more likely to get nurse visit cooperation in BHPSW19 and 11% more likely in USW2 compared to those under 55 years of age. For BHPSW19, nurses with over 16 years of experience are 14% more likely to get consent and collect blood compared to those with between 1 and 5 years of experience.

Similarly, for ELSA, age and experience of nurses are important. Older nurses are more likely to get blood consent in wave 2. Those 64 years and older have 10% higher chances to get consent compared to those under 55. Nurses with more experience are more likely to get nurse visit cooperation and collect blood in wave 6. Having more than 15 years of experiences increases the chances of collecting blood by 7% compared to those under 5 years of experience. Nurse ethnicity has no influence on participation.¹

Impact of nurses on population estimates

Lastly, we investigate how nurse characteristics influence weighted population estimates of biological markers. This is performed on estimates of mean pulse rate and CRP using four different weights that build on each other:

- Interview weights and area clustering
- + individual characteristics from the nurse data collection model
- + nurse characteristics from the nurse data collection model
- + nurse clustering

Comparing these different sets of weights can inform survey organizations whether nurse characteristics and clustering information are likely to impact substantive estimates of biological measures derived from the survey and nurse visit. Additionally, we investigate the difference in mean pulse rate and CRP by gender and long-standing illness for each weighting approach and survey. This is done to understand how group comparisons might be influenced by the specific weighting approach used.

Figure 3 presents the overall means with confidence intervals by outcome of interest, weight type, and survey. In all four cases including the nurse visit weights changes the mean estimate although this is statistically significant only in the case of CRP in ELSAW2. In all cases the nurse visit weights

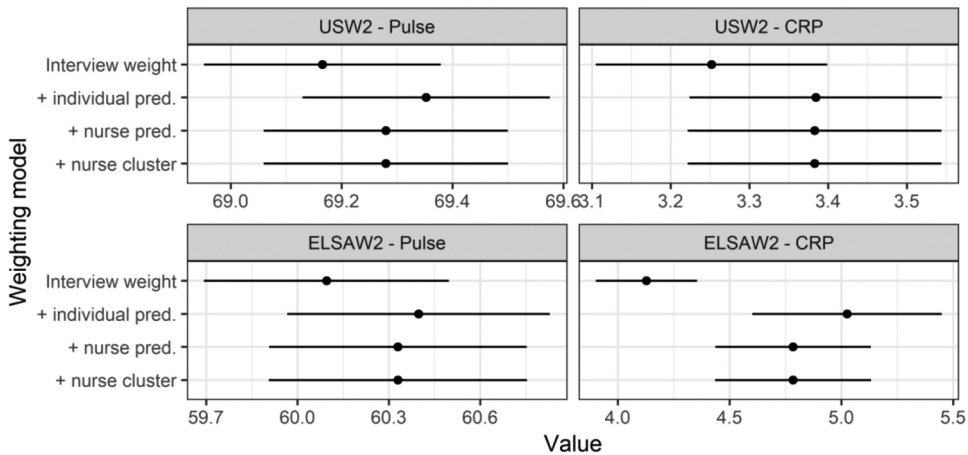


Figure 3. Pulse and CRP means with confidence intervals by survey and weighting approach.

lead to higher values for pulse rate and CRP, indicating worse health. Adding nurse characteristics (using model 2 versus model 1 to construct the nurse visit weights) shifts the means for three out of the four statistics although none of them are statistically significant. In all cases, including nurse characteristics decreases the mean, indicating a healthier overall population. Finally, adding nurse clustering (in addition to area clustering) does not change the statistic or the confidence interval.

Looking at the average pulse rate (Figure A4) and CRP (Figure A5) by gender and having a long-term illness leads to similar conclusions. The population estimates shift towards less healthy when the missing data are modeled in the nurse visit and this is slightly attenuated when including nurse characteristics. Including nurse clustering does not impact the statistics or the confidence intervals. In none of the cases do the conclusions change regarding group differences when using the new weights.

5. Conclusions and discussion

This paper investigated the effects of nurses on biomeasure participation in large longitudinal studies in the UK. We distinguished three non-response stages: nurse visit, consent to blood, and actual blood collection. We also looked at five different rounds of data collection. Two of them are general population surveys, one in wave 2 of UKHLS, and one in wave 19 of BHPS. The other three rounds come from waves 2, 4, and 6 of the ELSA study where the population of interest is above 50 years of age.

Across the five data collections, some consistent findings emerged. Firstly, nurses have an effect on non-response. This effect varies by stage but is present in all stages of the biomeasure component. As expected, nurse effects are smaller in the initial nurse visit stage, where nurses explain around 5% of variation. The nurse effect increases to around 9% for obtaining consent to collect blood while it reaches a maximum of 14% for actual blood collection. We consider these effect sizes to be low-to-medium and on par with known interviewer effects on biomeasures (Korbmacher, 2014; Sakshaug et al., 2010).

Regarding individual characteristics, we found some consistent patterns. Older people are more likely to participate in the nurse visit but the effects on consent and actual blood collection are mixed. Similarly, having no qualifications increases the chances of non-response to the nurse visit. Females are in general more likely to participate in the nurse visit but less likely to actually give blood. White respondents seem more likely to participate in all stages of data collection. People who own their house are more likely to participate in the nurse visit compared to those that don't. Respondents with better self-rated health are more likely to participate in all stages. Respondents with long-term health issues are more likely to participate in the nurse visit. Persistent patterns in the nurse effects showed that older nurses are more likely to gain cooperation to the nurse visit, those older than 64 years have

a higher chance to get blood consent compared to those under 55 years, and those who have over 16 years of experience are more likely to collect blood compared to those under 5 years of experience.

While we believe these findings bring important insights regarding non-response and participation in survey-based biological data collections, there are some limitations. The main limitation is how the nurse effects were estimated. Due to the lack of an experimental design randomizing nurses to respondents, we have relied on statistical models to estimate the nurse effects – a common strategy in interviewer studies (West & Blom, 2017). More precisely, cross-classified models were used controlling for area effects and individual and nurse characteristics. While the approach is imperfect, it is more practical than implementing a costly interpenetrated design. A second limitation is the limited amount of nurse information, so correcting for nurse effects and informing future survey designs are potentially limited. Thirdly, we control for item missing data only by including categories for the missing cases, which is often a contested practice (Groenwold et al., 2012). Finally, our investigation of non-response adjustment methods could be expanded by looking at different models for dealing with missing data (e.g. multiple imputation) and investigating different types of statistics (e.g. regression coefficients), although their performance will depend on the amount of auxiliary information available.

Despite these limitations, some general lessons can be extracted from this investigation. Firstly, nurses are important. They can explain a medium amount of variation in non-response, up to about 14%, and this is relatively consistent over multiple rounds of data collection. Secondly, by looking at the impact of nurse characteristics when modeling non-response and making inferences about the population, we showed that nurse characteristics have an impact. Specifically, including nurse characteristics in the weighting approach shifted the means for important health statistics. Although these shifts did not reach levels of statistical significance, one can imagine cases where borderline effects (e.g. differences between groups are small) could be changed by including nurse characteristics in the non-response model. Based on this finding, together with the significant nurse variation and nurse predictors, we recommend that nurse ID and nurse information be publicly released in all surveys and should be consistently used in models for non-response and post-survey adjustments (e.g. weighting). This is not currently the case for ELSA and UKHLS (S. McFall et al., 2014; Scholes et al., 2008). The results also highlight some individual respondent characteristics that should always be used in models for dealing with non-response in bio-social surveys. The key variables are gender, age, race, self-rated health, and long-term illness. These are essential as they are related to substantive research questions that are regularly investigated with these types of data.

Bio-social surveys have the great advantage of being based on probability samples. And while response rates are not 100% the ability to track and correct for the different stages of non-response is a great advantage compared to non-probability studies. As such, we do not believe our work undermines their high value. That being said, not all data providers are currently transparent about the impact of nurses on their results by making nurse data freely available to researchers. Having more data and research on this topic could improve our ability to minimize these nurse effects by design and correct for them post-data collection.

Future research should further investigate the impact of nurses in other surveys and countries to validate the findings reported here. Other potential avenues of research include the development of a nurse survey that could enhance our understanding of how and why nurses influence health data collections and an in-depth analysis of how the collected nurse characteristics can affect non-response adjustments.

Note

1. We ran a sensitivity analysis (results not shown) by including interactions in all models between the ages of the nurses and the ages of the respondents as well as between the ethnicities of the respondents and nurses. Out of approximately 100 new coefficients estimated only four were statistically significant. As such, we do not believe there is strong evidence of such interactions in the data.

Acknowledgments

We acknowledge the support of the National Centre for Research Methods and NatCen with acquiring the ELSA nurse data.

Disclosure statement

No potential conflict of interest was reported by the author.

References

- Al Baghal, T. (2016). Obtaining data linkage consent for children: Factors influencing outcomes and potential biases. *International Journal of Social Research Methodology*, 19(6), 623–643. <https://doi.org/10.1080/13645579.2015.1064635>
- Anglewicz, P. (2009). The malawi diffusion and ideational change project 2004–06: Data collection, data quality, and analysis of attrition. *Demographic Research*, 20(21), 503. <https://doi.org/10.4054/DemRes.2009.20.21>
- Banks, J., Breeze, E., Lessof, C., & Nazroo, J. (2006). *Retirement, health and relationships of the older population in England: The 2004 english longitudinal study of ageing (Wave 2)*. The Institute for Fiscal Studies.
- Blake, M., Marmot, M., Phelps, A., Steptoe, A., Oldfield, Z., Clemens, S., & Rogers, N. (2018). *English longitudinal study of ageing: Waves 0-8, 1998-2017*. UK Data Service.
- Bridges, S., Hussey, D., & Blake, M. (2015). *The dynamics of ageing: The 2012 english longitudinal study of ageing (Wave 6)* (pp. 1–100) [Technical Report]. NatCen.
- Cernat, A., & Sakshaug, J. W. (2020). The impact of nurse continuity on biosocial survey participation. *Survey Methods: Insights from the Field*, 1–14. <https://doi.org/10.13094/SMIF-2020-00010>
- Cheshire, H., Hussey, D., Medina, J., Piekering, K., Wood, N., Ward, K., & Lessof, C. (2012). *Financial circumstances, health and well-being of the older population in England: The 2008 English longitudinal study of ageing* (pp. 1–76). Technical series. The Institute for Fiscal Studies.
- Cochrane, A., Chapman, P. J., & Oldham, P. D. (1951). Observer errors in taking medical histories. *The Lancet*, 1 (6662), 1007–1009. [https://doi.org/10.1016/S0140-6736\(51\)92518-4](https://doi.org/10.1016/S0140-6736(51)92518-4)
- Cohen, G., & Duffy, J. C. (2002). Are nonrespondents to health surveys less healthy than respondents? *Journal of Official Statistics*, 18(1), 13–23. <https://www.scb.se/contentassets/f6bcee6f397c4fd68db6452fc9643e68/are-nonrespondents-to-health-surveys-less-healthy-than-respondents.pdf>
- Couper, M. P. (1997). Survey introductions and data quality. *Public Opinion Quarterly*, 61(2), 317–338. <https://doi.org/10.1086/297797>
- Durbin, J., & Stuart, A. (1951). Differences in response rates of experienced and inexperienced interviewers. *Journal of the Royal Statistical Society (Series: A)*, 114(2), 163–205. <https://doi.org/10.2307/2980975>
- Elosua, R., Bartali, B., Ordovas, J. M., Corsi, A. M., Lauretani, F., & Ferrucci, L. (2005). Association between physical activity, physical performance, and inflammatory biomarkers in an elderly population: The InCHIANTI study. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, 60(6), 760–767. <https://doi.org/10.1093/gerona/60.6.760>
- Groenwold, R. H. H., White, I. R., Donders, A. R. T., Carpenter, J. R., Altman, D. G., & Moons, G. M. (2012). Missing covariate data in clinical research: When and when not to use the missing-indicator method for analysis. *Research Methodology Series, Canadian Methodology Association Journal*, 184(11), 1265–1269. <https://doi.org/10.1503/cmaj.110977>
- Jaszczak, A., Lundeen, K., & Smith, S. (2009). Using nonmedically trained interviewers to collect biomeasures in a national in-home survey. *Field Methods*, 21(1), 26–48. <https://doi.org/10.1177/1525822X08323988>
- Kearney, P. M., Cronin, H., O'Regan, C., Kamiya, Y., Whelan, B. J., & Kenny, R. A. (2011). Comparison of centre and home-based health assessments: Early experience from the Irish longitudinal study of ageing (TILDA). *Age and Ageing*, 40(1), 85–90. <https://doi.org/10.1093/ageing/afq124>
- Kenny, R. A., Whelan, B. J., Cronin, H., Kamiya, Y., Kearney, P., O'Regan, C., & Ziegel, M. (2010). *The Design of the Irish Longitudinal Study on Ageing*. Trinity College Dublin.
- Knies, G. (2018). *Understanding Society: Waves 1-7, 2009-2016 and harmonised BHPS: Waves 1-18, 1991-2009, User Guide* (pp. 1–153). Institute for Social and Economic Research.
- Korbmacher, J. M. (2014). *Interviewer effects on respondents' willingness to provide blood samples in SHARE* (Working Paper Series 20-2014). Munich: Survey of Health, Ageing and Retirement in Europe (SHARE).
- Lindau, S. T., Hoffmann, J. N., Lundeen, K., Jaszczak, A., McClintock, M. K., & Jordan, J. A. (2009). Vaginal self-swab specimen collection in a home-based survey of older women: Methods and applications. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 64(suppl_1), i106–i118. <https://doi.org/10.1093/geronb/gbn021>

- Lumley, T. (2004). Analysis of complex survey samples. *Journal of Statistical Software*, 9(8), 8. <https://doi.org/10.18637/jss.v009.i08>
- McFall, S., Petersen, J., Kaminska, O., & Lynn, P. (2014). *Waves 2 and 3 nurse health assessment, 2010-2012. Guide to nurse health assessment* (Understanding Society: The UK Household Longitudinal Study) (pp. 1–38). Institute for Social and Economic Research. https://www.understandingsociety.ac.uk/d/100/7251_User_Guide_Health_Assmt_w2_w3.pdf?1392855567
- McFall, S. L., Conolly, A., & Burton, J. (2014). Collecting biomarkers using trained interviewers. Lessons learned from a pilot study. *Survey Research Methods*, 8(1), 57–66. <https://doi.org/10.18148/srm/2014.v8i1.5471>
- Mostafa, T. (2016). Variation within households in consent to link survey data to administrative records: Evidence from the UK Millennium cohort study. *International Journal of Social Research Methodology*, 19(3), 355–375. <https://doi.org/10.1080/13645579.2015.1019264>
- Mostafa, T., & Wiggins, R. D. (2018). What influences respondents to behave consistently when asked to consent to health record linkage on repeat occasions? *International Journal of Social Research Methodology*, 21(1), 119–134. <https://doi.org/10.1080/13645579.2017.1332848>
- Muthén, L. K., & Muthén, B. O. (2017). *Mplus user's guide*, (8th ed.). Muthén & Muthén.
- NatCen Social Research. (2015). English longitudinal study of ageing (ELSA) Wave 2, 4 and 6 user guide to the nurse datasets (pp. 1–28).
- National Research Council. (2001). *Cells and surveys: Should biological measures be included in social science research?* The National Academies Press.
- National Research Council. (2008). *Biosocial surveys*. The National Academies Press.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41–55. <https://doi.org/10.1093/biomet/70.1.41>
- Sakshaug, J. W., Couper, M. P., & Ofstedal, M. B. (2010). Characteristics of physical measurement consent in a population-based survey of older adults. *Medical Care*, 48(1), 64–71. <https://doi.org/10.1097/MLR.0b013e3181adcbd3>
- Sakshaug, J. W., Ofstedal, M. B., Guyer, H., & Beebe, T. J. (2015). The collection of biospecimens in health surveys. In T. Johnson (Ed.), *Handbook of health survey methods* (pp. 383–419). John Wiley and Sons.
- Scholes, S., Taylor, R., Cheshire, H., Cox, K., & Lessof, C. (2008). Retirement, health and relationships of the older population in England: The 2004 English longitudinal study of ageing (pp. 1–101). The Institute for Fiscal Studies.
- Seeman, T. E., Huang, M. H., Bretsky, P., Crimmins, E., Launer, L., & Guralnik, J. M. (2005). Education and APOE-e4 in longitudinal cognitive decline: MacArthur studies of successful aging. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 60(2), P74–P83. <https://doi.org/10.1093/geronb/60.2.P74>
- Snijders, T. A. B., & Bosker, R. (2011). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. Sage Publications Ltd.
- Steinkamp, S. W. (1966). Some characteristics of effective interviewers. *Journal of Applied Psychology*, 50(6), 487–492. <https://doi.org/10.1037/h0023998>
- Taylor, R., Conway, L., Calderwood, L., Lessof, C., Cheshire, H., Cox, K., & Scholes, S. (2007). Health, wealth and lifestyles of the older population in England: The 2002 English longitudinal study of ageing (pp. 1–74). dcxHealth, wealth and lifestyles of the older population in England: The 2002 English longitudinal study of ageing.
- Tolonen, H., Helakorpi, S., Talala, K., Helasoja, V., Martelin, T., & Prättälä, R. (2006). 25-year trends and socio-demographic differences in response rates: Finnish adult health behaviour survey. *European Journal of Epidemiology*, 21(6), 409–415. <https://doi.org/10.1007/s10654-006-9019-8>
- University of Essex. (2014). *Understanding society: Waves 2-3 nurse health assessment, 2010-2012* (Data set). UK Data Service. <https://doi.org/10.5255/UKDA-SN-7251-3>
- University of Essex. (2018). *Understanding society: Waves 1-7, 2009-2016 and Harmonised BHPS: Waves 1-18, 1991-2009*. UK Data Service.
- Weiss, L. M., Sakshaug, J. W., & Börsch-Supan, A. (2019). Collection of biomeasures in a cross-national setting: Experiences in SHARE. In T. Johnson, B.-E. Pennell, I. Stoop, & B. Dorer (Eds.), *Advances in comparative survey methods: Multinational, multiregional and multicultural contexts* (3MC) (pp. 623–641). John Wiley and Sons.
- West, B. T., & Blom, A. G. (2017). Explaining interviewer effects: A research synthesis. *Journal of Survey Statistics and Methodology*, 5(2), 175–211. <https://doi.org/10.1093/jssam/smw024>
- Williams, S. R., & McDade, T. W. (2009). The use of dried blood spot sampling in the national social life, health, and aging project. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 64(suppl_1), i131–i136. <https://doi.org/10.1093/geronb/gbn022>
- Zipf, G., Chiappa, M., Porter, K. S., Ostchega, Y., Lewis, B. G., & Dostal, J., . (2013). National health and nutrition examination survey: Plan and operations, 1999–2010. National center for health statistics. *Vital Health Statistics*, 1, 56, 1–37. <https://pubmed.ncbi.nlm.nih.gov/25078429/>